

Transparent Active Learning for Robots

Crystal Chao, Maya Cakmak, Andrea L. Thomaz

School of Interactive Computing

Georgia Institute of Technology

Atlanta, Georgia 30332, USA

{cchao, maya, athomaz}@cc.gatech.edu

Abstract—This research aims to enable robots to learn from human teachers. Motivated by human social learning, we believe that a transparent learning process can help guide the human teacher to provide the most informative instruction. We believe active learning is an inherently transparent machine learning approach because the learner formulates queries to the oracle that reveal information about areas of uncertainty in the underlying model. In this work, we implement active learning on the Simon robot in the form of nonverbal gestures that query a human teacher about a demonstration within the context of a social dialogue. Our preliminary pilot study data show potential for transparency through active learning to improve the accuracy and efficiency of the teaching process. However, our data also seem to indicate possible undesirable effects from the human teacher’s perspective regarding balance of the interaction. These preliminary results argue for control strategies that balance leading and following during a social learning interaction.

Index Terms—active learning, socially guided machine learning, social robots, interactive learning, human-robot interaction

I. INTRODUCTION

In many of the scenarios that people envision for personal robots (e.g., general assistants in homes, schools, and hospitals), it is hard to imagine how a robot designer can possibly pre-program a robot with all of the skills necessary for that domain. Our research is about developing ways for these robots to learn the necessary tasks and skills from end users—Socially Guided Machine Learning (SG-ML). Because we cannot expect these users to have extensive experience with machine learning or robotics, we need to design algorithms and systems that take advantage of the ways that they naturally approach the task of teaching.

In the work presented here, we focus on a particular aspect of SG-ML: how a robot learner can improve its own learning environment through transparency. When situated in an interaction, a “black box” learning process does nothing to improve the quality and relevance of the instruction it receives. Transparency of the internal state of the machine could improve the learning experience by revealing to the teacher what is known and what is unclear. By guiding the teaching process, the machine can learn more quickly as well as reduce the workload for the teacher.

In this paper, we look to active learning as a method of achieving transparency. Active learning refers to machine learners that can select the examples to be labeled by an oracle, potentially reducing the number of examples required. In SG-ML, the human teacher serves as the oracle. We analyze how active learning can be integrated with nonverbal social cues

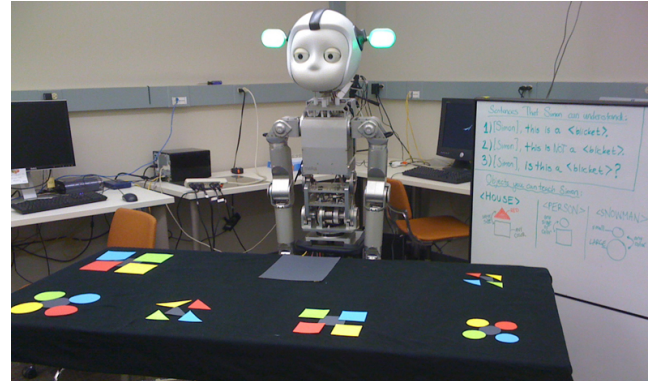


Fig. 1. The Simon robot’s workspace from the human teacher’s perspective.

to make ambiguity in the underlying model transparent by having the robot solicit appropriate examples from a human partner in an intuitive way. We present our active learning implementation, a design for an experiment with human teachers, and initial observations from a pilot study of that design. Preliminary results highlight the method’s potential to improve performance but also argue for a more balanced control of the interaction than traditional active learning.

II. APPROACH

In aiming to build more teachable robots, child development and the human learning process serve as inspiration and motivation. In particular, situated learning is a field of study that looks at the social world of a child and how it contributes to the child’s development [1]. In a situated learning interaction, a good instructor maintains a mental model of the learner’s understanding and structures the learning task appropriately with timely feedback and guidance. The learner contributes to the process by expressing internal state via communicative acts (e.g. expressing understanding, confusion, attention, etc.) [2], [3]. This reciprocal and tightly coupled interaction enables the learner to leverage the teacher’s instructions to build the appropriate representations and associations.

This human-style tutelage is a cooperative and social activity. Theories of human collaboration argue for an open line of communication between teacher and learner in order to best maintain a mutual belief about the learner’s knowledge state [4], [5]. These theories inform the design of SG-ML systems.

The scenario of human input has received some attention in the machine learning and robotics community. Much prior

work deals with the scenario where a machine learns by observing human behavior [6], [7], [8]. Other work has focused on how a machine can learn tasks from human instruction [9], [10], with human advice [11], [12], or through intervention in the learner's actions during the learning process [13]. Our work in transparency and active learning is complementary to many of these task learning approaches. Specifically, our goal is to have more *interactive* systems that can learn in real-time from everyday people.

Related to active learning in particular, there has been work on computational benefits of teacher/learner pairs [14]. Active learning is a field that addresses efficient use of an external oracle [15], [16], [17]. One system shows confidence-based active learning with (non-social) human labelers [18]. Queries can be viewed as a type of transparency into the learning process, which is why we are interested in applying active learning to the context of robot learning in human-robot interaction. The challenge is in designing ways that the robot can use natural social cues to solicit appropriate examples from a human partner in the learning task.

Active learning research focuses on how to select queries that maximize information gain or otherwise make efficient use of the oracle. In contrast, this work is concerned with using active learning for robot task learning, as well as the challenges of allowing the oracle to be a non-expert in machine learning. In our experiments, we evaluate performance of the machine learner when taught by human teachers and ask the question: can the robot influence the teacher to provide the examples that it wants? And how does this affect the ease and speed of teaching from the human partner's perspective?

In this paper we present a learning system, an experiment with human teachers, and observations drawn from this data.

III. PLATFORM

The robotic platform for this research is "Simon," an upper-torso humanoid social robot with 7-DOF arms, 4-DOF hands, and a socially expressive head and neck (Fig. 1). We are developing the Simon platform specifically for face-to-face human-robot interaction. In our task scenarios, the robot works alongside or across from a human partner at a tabletop workspace. The robot has the ability to perform simple gestures (e.g., pointing, head nods and shakes) to communicate about objects that the human can use for teaching.

Our learning system is implemented within the C6 software system (see [19]), which has a specific pipeline for triggering robot actuations from sensory inputs. A robot is equipped with various sensors such as cameras for vision and microphones for speech recognition. At every time step, the robot receives observations $O = \{o_1, \dots, o_k\}$ from these sensory processes. The perception system is a set of percepts $P = \{p_1, \dots, p_n\}$. Each $p \in P$ is an atomic classification and data-extraction unit that models an aspect of each observation from the sensory system by returning a match probability such that $p(o) = m$, where $m \in [0, 1]$ is a match value. The percept provides a useful level of abstraction for reducing the dimensionality of incoming sensory information.

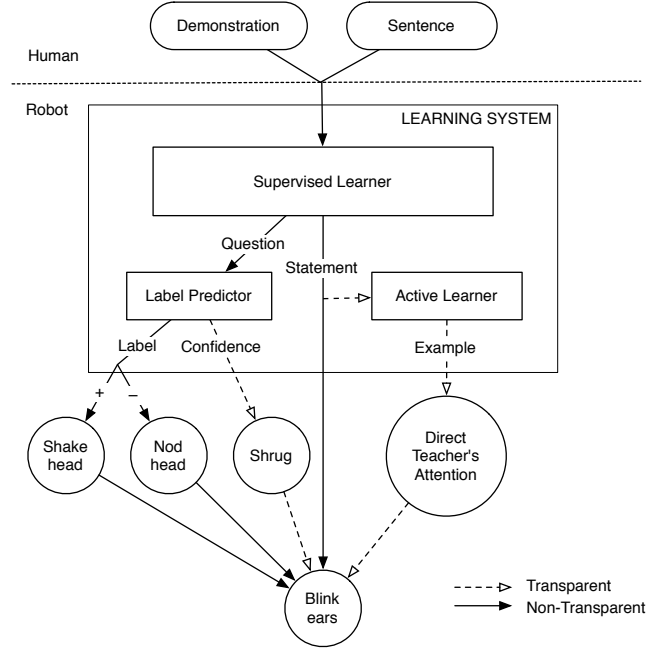


Fig. 2. System diagram of the learning and interaction.

The belief system maintains the belief set B by integrating these percepts into discrete object representations (based on spatial relationships and various other similarity metrics or tracking mechanisms). Belief objects that detail the perceived state of the world are sent to the action system for decision-making. The action system is structured as action groups of hierarchical action tuples requiring preconditions, executables, and postconditions. After a high-level action is selected, the lower-level joint trajectories are rendered in the motor system to the virtual character in simulation as well as transmitted to the motor module controlling the physical robot.

In this paper we describe extensions to the C6 system that enable transparent active learning.

IV. LEARNING ENVIRONMENT

A. Domain Description

In this work, the task domain involves colorful paper cutouts, which we somewhat imprecisely call *tangrams*. Simon's learning task involves recognition of paired configurations of these tangrams. We believe this domain is interesting because it is related to the general class of assembly tasks.

Each tangram has a *color* attribute with four possible values (red, green, yellow, or blue), a *shape* attribute with three possible values (square, triangle, or circle), and a *size* attribute with two possible values (small or large), for a total of 24 possible unique tangrams. Each configuration consists of a top and a bottom tangram and has additional meta-attributes of whether the attributes for the top and bottom tangrams have the same or different values. We call these configurations *compound tangrams*. Thus, each compound tangram has a total of 9 attributes: the 3 attributes for the top, the 3 attributes

for the bottom, and the 3 meta-attributes for whether the top and bottom match. Simon’s workspace contains exactly one of each tangram, so there are 552 possible compound tangrams in his workspace.

B. Perception for Domain

Tangrams lying on a table in front of Simon are detected through a fixed overhead camera and segmented using background subtraction. The shape of the tangram is recognized by the number of corners of the simplified polygon contour of the segmented object (square: 4, circle: 8, triangle: 3). Size is recognized based on the area within the contour and color is recognized using the color histogram of the segmented object. The tangrams are localized in robot world coordinates using a fixed homography from the image plane to the table plane.

C. Simon’s Workspace

Simon’s workspace is a table covered by a black tablecloth. The 24 possible tangrams are arranged in 6 separate groups on the perimeter of this table. Each group contains tangrams of the same size and shape, but of varying color. The space immediately in front of Simon is the demonstration area for positive and negative examples; its position is demarcated to the human with a black sheet of paper (Figure 1).

The human teacher stands on the opposite side of the table from Simon for face-to-face interaction and provides examples in the demonstration area from Simon’s perspective. A whiteboard next to Simon provides reminders to the human about the experiment instructions.

V. LEARNING SYSTEM

This section describes our implementation of supervised learning for symbol generalization and active learning for example selection.

A. Symbol Representation

A symbol definition is a conjunctive subset of the possible compound tangram attribute values. This subset describes the attribute values that must hold true in order for a compound tangram to be a member of that symbol class. For example, one of the symbols we use is a small circle on top and a large circle on bottom, which we call “snowman.” In this paper, Simon is tasked with learning four different symbols from demonstrations provided by human teachers.

Any given example of a compound tangram can be described using a version space, which is the set of all symbol hypotheses consistent with that example [20]. A hypothesis that is consistent with an example is one that correctly classifies that example.

B. Generalization Through Supervised Learning

1) *Symbol Version Space*: Simon can generalize the correct symbol definition by seeing demonstrated examples with positive and negative labels. The generalization is performed using the following supervised learning mechanism.

For each positive demonstration of a compound tangram, all of the hypotheses in the version space for that demonstration

TABLE I
SIMPLIFIED EXAMPLE: SYMBOL GENERALIZATION OF “Red”

Step	Example	Label	Version Space
1	Small red circle	+	Red, Small, Circle Red, Small Red, Circle Small, Circle Red Small Circle
2	Small green circle	−	Red, Small, Circle Red, Small Red, Circle Red
3	Large red circle	+	Red, Circle Red
4	Small red square	+	Red

are expanded. Redundant hypotheses due to meta-attributes are filtered out. The remaining hypotheses are tracked for consistency throughout the teaching process for that symbol.

Tracked hypotheses are also ranked. We quantify a hypothesis’s *consistency* as the number of examples from the demonstration data that are consistent with that hypothesis, and then rank the tracked hypotheses using this metric. The version space for the symbol thus becomes the set of hypotheses that are of the highest consistency value among all of the tracked hypotheses.

For each provided negative example, we do not add any new hypotheses to track, but we adjust the consistency ranking of the existing hypotheses to accommodate the new label.

2) *Example Generalization*: To illustrate the generalization, we provide an example in Table I using the simplified domain of single tangrams to learn the concept of “red.” This domain contains only the three attributes of color, shape, and size. Over four example demonstrations, the learner generalizes such that the best hypothesis is the one in which only color matters, and the color attribute has a value of “red.”

C. Example Selection Through Active Learning

Active learners have a mechanism for selecting examples to be labeled by an oracle [17]. Ideally, these examples are maximally informative to the learner and thus significantly reduce the number of required labels to create a good model.

For Simon’s active learning mechanism, we implement query-by-committee for example selection [21]. This method uses a committee of competing hypotheses and selects the example that results in maximal disagreement between the hypotheses in terms of predicted labels. The committee in our implementation is the symbol version space — that is, the set of all maximally ranked hypotheses being tracked while learning the symbol. We assign the value of +1 for a positive label and −1 for a negative label. This means that the example with the smallest absolute-valued average predicted label value over the entire symbol version space will result in the highest information gain. The effect of iteratively labeling examples selected by the committee is to prune away as much of the

committee as possible until only one correct hypothesis is left, reducing the entropy of the committee.

Note that there are often many examples with the same best value; in the example in Table I, there are actually 6 best examples after step 1, 12 best examples after step 2, and 16 best examples for step 3. In these cases, the active learner can impose some additional ordering or select randomly. We discuss additional constraints used by the active learner in the next section on transparency mechanisms.

With a committee of hypotheses, label prediction occurs by using the majority label from the committee, which is also the sign of the average label. Our measure of confidence in the label is the distance of the majority label to the average label. Thus for the degenerate case of having no majority label, the confidence is 0.

VI. TRANSPARENCY MECHANISMS

We implemented two types of transparency mechanisms for this experiment. One focused on communicating points of confusion for the robot as determined by the active learning component. The other focused on communicating the robot’s confidence or uncertainty to the human.

A. Attention Direction

When performing active learning queries of the hypothesis space, we would ideally like Simon to pick up relevant objects and replace them in the demonstration to solicit a label from the teacher. Alternatively, Simon could simply talk about the objects. While developing the speech synthesis and manipulation capabilities for such an experiment, we are doing preliminary studies with a more indirect route for communicating a query.

We limit example selection to compound tangrams that share either the top or the bottom tangram with the current demonstrated compound tangram. This allows for a reduction in the complexity of communicating the example, since the robot only needs to select a single tangram to replace in the current demonstration. In order to draw the human teacher’s attention to this desired replacement, the robot first points at either the top or the bottom tangram to replace. It then gazes at one of the six groups of tangrams in the workspace, to the group with the appropriate size and shape. Finally it lights up its ears in the color of the desired example. For example, pointing at the bottom half of the demonstration area while gazing with red ears at the small circles is meant to ask, “Can we put the small red circle here?”

B. Test Responses

When the human teacher tests the robot on an example, Simon has a total of five possible responses: yes, uncertain-yes, uncertain, uncertain-no, and no. These are communicated using three animations, one each for “yes,” “no,” and “uncertain.” The “yes” animation is the head nodding, and the “no” animation is the head shaking from side to side. The “uncertain” animation is designed to communicate lack of confidence. It tilts the head sideways and lifts the arms upwards in a

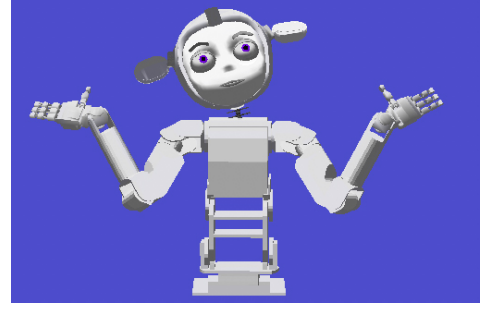


Fig. 3. The “uncertain” shrugging animation in simulation.

TABLE II
SYMBOL LEARNING TASKS

Name	Attributes	Values	Generality
House	Top shape Top color Bottom shape Match size	Triangle Red Square Same	1.45%
Snowman	Top shape Top size Bottom shape Bottom size	Circle Small Circle Large	2.89%
Person	Top shape Bottom shape	Circle Square	10.87%
Rainbow	Match color	Different	75.00%

bent-elbow configuration to create a shrugging gesture (Figure 3). By playing the “uncertain” animation simultaneously with either “yes” or “no,” Simon’s response can communicate that he is leaning towards one answer without being fully confident in that answer.

VII. EXPERIMENT DESIGN

A. Teaching Task

Subjects were tasked with teaching Simon four different symbols. We named these symbols “house,” “person,” “snowman,” and “rainbow.” The ground truth representation for each symbol is shown in Table II, and examples can be seen in Fig. 4. We tried to select these symbols based on varying specificity/generality — that is, these symbols have varying coverage of the space of possible compound tangrams.

We described the feature space to the subjects and told them to teach the symbols one at a time. Subjects were allowed to teach “house,” “person,” and “snowman” in any order, but “rainbow” was always last. We wanted subjects to become familiar with the objects before trying to teach a more general concept.

In order to teach Simon, subjects were told to arrange a compound tangram in the demonstration area from Simon’s perspective and say one of three possible sentence types into a microphone:

- [Simon], this is a <symbol-name>.
- [Simon], this is *not* a <symbol-name>.
- [Simon], is this a <symbol-name>?

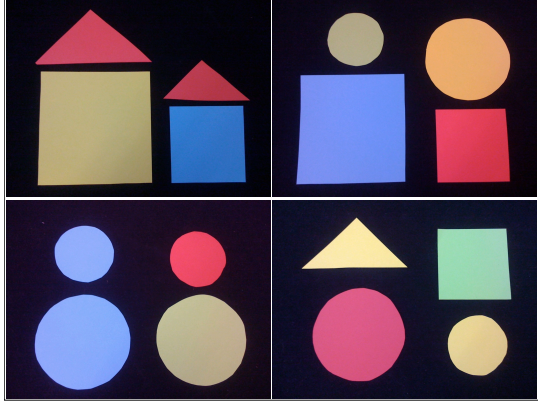


Fig. 4. Examples of the four symbols that people taught Simon: House (top-left), Person (top-right), Snowman (bottom-left), Rainbow (bottom-right).

These sentences were originally to be identified using a Sphinx speech recognition module, but were ultimately logged by the experimenter using a Wizard of Oz interface for higher accuracy. The subjects still spoke into a microphone to preserve the belief of communicating with the robot.

Taken in conjunction with an example compound tangram, these three sentences represented a positive label, a negative label, and a test question. Simon would then process this statement and respond to it. Subjects were instructed to watch for Simon’s reaction, since he could not speak. Simon blinked the lights on his ears when he was finished responding to indicate to the teacher that he was ready for another demonstration. Subjects were told to put any pieces used for demonstrations back in their original locations in the workspace when not currently in use.

A whiteboard near Simon provided reminders about the symbol attributes and the types of sentences that the teacher could say. This was preferred over giving the human a piece of paper with instructions so that people would look up towards the robot rather than fixate on the piece of paper.

Subjects were instructed to continue teaching each symbol until they were satisfied that Simon had learned the symbol well or thought that Simon had stopped making progress.






B. Experimental Conditions

In this experiment we wanted to compare transparent active learning to a more typical supervised learning interaction. We used a between-subjects design comparing two groups:

1) *Transparent Group (T)*: All of the transparency mechanisms we implemented were included in interactions for the transparent group. An example interaction sequence is shown in Table III. After each positive or negative label provided, the robot looked at the demonstration, processed it using supervised learning, and then attempted to direct the human’s attention towards an example selected by active learning using the multimodal gesture of pointing, gazing, and ear color.

After each test question, the robot responded in one of five

TABLE III
TRANSPARENT GROUP EXAMPLE INTERACTION

Step	Demo Area	Teacher’s Sentence	Robot’s Response
1		“This is a house.”	
2		“This is NOT a house.”	
3		“Is this a house?”	(Shrug gesture)

possible ways: certain-yes, uncertain-yes, uncertain, uncertain-no, or certain-no. Certain responses were predictions of confidence value 1; confidences lower than this resulted in playing back the shrugging animation. Yes or no responses were the nodding and head-shaking animations respectively. Uncertain-yes and uncertain-no were thus combinations of simultaneous head and body animations.

When instructing the transparency group subjects, we said that the robot knew about the top and bottom of the demonstration area as well as the groups of shapes in his workspace. We emphasized that the robot may try to communicate with them about the example provided, but did not describe any gestures to expect.

2) *Non-Transparent Group (NT)*: In the non-transparent group, Simon did not perform active learning. He also never played back the shrugging animation to indicate uncertainty. Thus, the only possible responses were the nodding and head-shaking animations. Like the transparent group, we did not tell these subjects what gestures to expect.

C. Data Logged

We considered an important event to be either a teacher’s sentence or the robot’s response. The following were the data about each event that we logged for every symbol teaching session in order to characterize the differences between the transparent and non-transparent groups:

- *Transparent* – whether this was the T or NT group.
- *System Time* – wall clock time.
- *Interaction Step* – number of events so far for this symbol.
- *Current Example* – the compound tangram currently configured in the demonstration area.
- *Current Label* – the label provided for the compound tangram, if any.
- *Sentence Type* – one of the three valid sentences types from the human teacher, if any.
- *Answer Type* – one of the five responses the robot could provide, if any.
- *Query* – the compound tangram selected for attention direction using active learning, if any.

D. Interview Questions

After each subject finished teaching all of the symbols, the experimenter asked the subject a series of questions to evaluate the teaching experience.

- What was Simon communicating with his ear color, gaze, and pointing? His head and body gestures?
- What was the difficulty of teaching each concept?
- How well do you think Simon learned each concept?
- What was your general teaching strategy?
- When did you decide to stop each teaching session?
- What else would you have liked to communicate?

VIII. PILOT RESULTS

We ran a pilot study on nine subjects in order to identify issues in the experiment design and witnessed a surprisingly diverse range of interactions. Because of the small sample size, none of the results here are statistically significant. However, we thought it would be interesting to present some salient characteristics of these interactions in a descriptive manner. These observations raise important issues to consider when designing a transparent robot learner.

A. Ambiguity in Transparency Mechanisms

One of the main conclusions we are able to draw from our pilot study is that it is important for the transparency devices to be clearly understandable. In these initial experiments, we found that 5 of the 7 people in the transparent group did not “correctly” interpret Simon’s nonverbal gestures as queries about the learning space. People had a variety of interpretations as to what the gestures specifically meant (such as pointing at a shape that was relevant or flashing red ear color when angry), but essentially the only information they understood from the robot’s communication was the fact that it needed more examples, not specifically what kind of examples were needed. This seems to have led to worse performance than the non-transparent group. Our hypothesis is that the robot’s gestures caused people to form a variety of different inaccurate mental models about the learning process that in turn led to inefficient teaching, whereas in the non-transparent case, people devised a teaching strategy to take the lead and systematically explore the feature space. In this semi-transparent group, ambiguous gestures seemed to prevent subjects from attempting this efficient systematic exploration.

This leads us to believe that we should wait to run the full version of this experiment until we have implemented either manipulation capabilities or speech synthesis. This way, the robot can use actions or speech as queries rather than relying on nonverbal gestures that were ambiguous to a majority of the subjects in the pilot study.

For the remainder of this discussion we will refer to three groups of people instead of two: the fully transparent group (FT) of two that understood the transparency mechanisms, the semi-transparent group (ST) of five that saw Simon’s transparency gestures but did not interpret them as intended, and the non-transparent (NT) of two that were never presented the transparency mechanisms.

TABLE IV
EFFICIENCY: ACCURACY DIVIDED BY THE TIME TAKEN TO TEACH

Group	Control	Accuracy	Time	Efficiency
(FT)	Single Initiative	93.75%	20.33 mins	18.79%/min
(NT)	Single Initiative	79.52%	24.56 mins	13.45%/min
(ST)	Mixed Initiative	70.76%	31.94 mins	10.74%/min

B. Performance Comparison

In order to compare the relative performance between the groups, we look at the *accuracy* of learned symbols, comparing the trained model’s predictions against the ground truth representation for the 552 possible compound tangrams. We used the F_1 score in our evaluations.

Of the two subjects in the FT group, one attempted to execute every example suggested by the robot and did not begin testing the robot at all until the robot stopped example selection. This subject taught models of 100% accuracy for all four symbols. The other subject did not recognize the active learning component until the third symbol learned, but his third and fourth symbols were 100% accurate. The ST and NT groups had an average accuracy of 70.76%.

Additionally, we can characterize the *efficiency* of teaching between the three groups. Efficiency is defined as the accuracy divided by the wall clock time taken to teach. As shown in Table IV, the FT group had the best efficiency, followed by NT, and then ST.

These performance outcomes suggest that it is possible for active learning to be extremely beneficial to the robot’s learning when being taught by a human. By trying to influence the human’s attention towards the areas of maximal confusion for the robot, the feature space for the hypotheses can be covered as completely and efficiently as possible and converge on perfect models rather quickly. In addition, subjects who did not participate in active learning remarked on how much mental work was required to keep track of different combinations in order to teach systematically. During active learning, the robot’s resources of memory and processing power can be leveraged instead of taxing the human teacher.

C. Perceived Performance

To analyze people’s mental model of the learning process, we compare actual accuracy to perceived accuracy. We asked people to estimate the percentage of each symbol that they thought Simon would be able to recognize in the future. The disparity between the subjects’ perception of accuracy and the accuracy of the actual models learned provides further support for an active learning approach.

The one subject who fully cooperated in active learning was confident that Simon had learned all of the symbols perfectly. For all other subjects there was no correlation between how well people believed Simon learned the concepts and how well Simon’s learned models actually performed. For example, some subjects estimated Simon’s learning a symbol as “50% complete” when the learned model was actually perfect. Other

subjects thought that certain symbols were learned significantly better than others when in fact the reverse was true.

Ideally, perception of accuracy should be matched to actual accuracy so that the teacher does not waste work or stop teaching prematurely. Because it can be challenging for humans to maintain a mental model of the learning space, active learning with unambiguous transparency mechanisms can potentially serve to shift work in the interaction from the human teacher to the robot learner.

Perceived accuracy is a metric that we would like to develop further in future experiments. In general, an important direction for our future work is recreating that quality of a sufficiently transparent process that leads to matched confidence between the teacher and the learner. In particular, we continue to look for appropriate metrics and methods for evaluating a person's mental model of the teaching process.

D. Single-Initiative Speed

Even disregarding accuracy, we observed an interesting outcome in time taken to complete the teaching tasks. The FT and NT subjects took less time to reach a point when they were satisfied with the robot's learned symbol. These subjects completed the study in an average of 22.44 mins (57.75 interaction steps, where a step is either a human sentence or a robot response), compared to the 31.94 mins (65.67 steps) of the ST group.

Our hypothesis is that "single-initiative" control (FT and NT groups) of the interaction resulted in faster learning compared to "mixed-initiative" control (ST group). In the FT group, the robot was in control of the learning process, with comparatively little input from the human teacher in terms of example selection. The robot completed learning quickly due to the active learning component. In the NT group, the human was in control of the robot's learning process because the human had the sole responsibility of example selection. Although the non-transparent teachers were provided less information, the result was that they concentrated on forming mental models of the symbol so that they could traverse the feature space systematically and teach the robot efficiently.

The accuracy of the ST group that understood the "uncertain" gestures but not the active learning gestures was similar to the NT group, but they took a longer time to achieve this. To analyze this difference, we looked at the percentage of demonstrations that were *uninformative*. We defined uninformative examples to be ones that did not change the size of the maximally consistent symbol version space at the time of presentation; these were examples that the robot already understood how to classify perfectly. The ST group provided an average of 29.67% uninformative demonstrations, whereas the NT provided an average of 15.63% uninformative examples. This could be explained by their impression that it wasn't necessary to keep track of examples provided because the robot seemed to be doing it, and then their attempts to reconcile conflicts between this impression and the reality of the robot's learned model. ST subjects kept providing examples when the robot demonstrated uncertainty after a test

question, but without having kept track of previous demonstrations, they provided these examples seemingly randomly and often redundantly. Additionally, when accuracy is taken into account, we also see that both FT and NT groups were more efficient than the ST group (see Table IV).

This highlights the importance of designing transparency mechanisms that convey accurate information at the appropriate level of abstraction. The semi-transparent group subjects reported feeling confusion at the active learning gestures. Although they saw more transparency mechanisms than the non-transparent group, this did not increase overall efficiency in teaching the task.

E. Balance of Control

Transparency through active learning has the potential to improve performance, but our preliminary data indicates that an interaction dictated entirely by active learning may be undesirable from the human teacher's perspective. Both of the subjects we saw in this condition expressed a desire to take a more active role in structuring the learning process rather than blindly answering all of the robot's questions.

The subject who cooperated fully with the robot's active learning was mostly reliant on the robot to guide the interaction. He stated that he was glad that the robot was able to figure out for itself what it needed to know, and that he stopped bothering to keep track of what the robot could be "caught on." Although he said he trusted the robot more as time went on, he also said that he was "skeptical" when the robot stopped asking questions and thus asked many test questions to the robot. On multiple occasions he also inserted an example that the robot did not request because he wanted to give demonstrations that he believed to be representative, just in case. The result was that he actually gave 12.19% uninformative examples; these instances were ultimately unnecessary. So although he felt compelled to take control of the interaction occasionally, these actually decreased the efficiency of the learning process.

The subject who participated in active learning for his last two symbols had a different situation. When he recognized the robot's active learning strategy, he considered it inefficient due to the constraint of varying only a single tangram; he thought that varying both the top and the bottom would be more efficient. Knowing the limitation of his own memory, he said he wished he could trust the robot's selected example but didn't trust it because of this perceived deficiency. He also said he felt like he wanted to "disobey" what the robot wanted to be taught, and that he sometimes "gave in" to the robot out of "pity." One ST subject, upon learning the purpose of the active learning gestures, exclaimed that she felt bad for Simon that she didn't answer any of his questions. These remarks suggest that active learning exerts a strong influence on the interaction in swaying the balance of control towards the robot, which could actually be an undesirable effect in an interaction.

On the opposite extreme were the NT subjects, during whose interactions the human had full control of the learning. In their exit interviews, these subjects explicitly stated either that they wished the robot could ask questions, or that

they wished they could ask the robot what exactly it didn't understand. In essence, these subjects felt that the robot did not participate enough in its own learning process and needed to maintain a higher amount of control in the interaction.

Thus, our hypothesis for the full version of this experiment is that people in both the transparent and non-transparent conditions will be dissatisfied with the balance of control in the interaction, but for different reasons. Moreover, we expect several people in the transparent condition to ignore the robot's queries from time to time in order to exert their own structure and control on the process.

We expect that this particular aspect can vary significantly due to differences in human teachers' personalities, so it will be necessary to have an adequate sample size to cover this variance in the final experiment. But we think it is important to state that balance of control between the robot and the human is a parameter to tune carefully when designing an interaction, as either extreme can have undesirable ramifications. Robots need a more fluid balance of control between leading and following. The open research question is how to design appropriate mechanisms for deciding how to switch between the two behaviors in a social learning interaction.

IX. CONCLUSION

In this work, we strove to build a robot learner that could leverage transparency in order to help a human teacher provide better instruction. We implemented active learning on the Simon robot so that it could query an external entity about areas of uncertainty in its hypothesis space. We then developed a set of nonverbal gestures for providing feedback about uncertainty, as well as nonverbal gestures that the robot could use to ask a human teacher about the feature space in a tangram symbol learning task.

In a pilot study, we saw that transparency through active learning has the potential to improve both the accuracy and the efficiency of a teaching process. We saw that people who understood the robot's queries were able to train a model with perfect accuracy relatively quickly. They also had high confidence that the trained model was complete. However, we decided that an interaction guided completely by active learning could be undesirable from the human teacher's perspective, since both of the subjects we saw in this condition indicated that they wanted to take more control of the process. Conversely, the subjects in the non-transparent condition who did not see the active learning version of the robot indicated that they wished that the robot would take some initiative or be able to communicate about what it did or did not know.

More broadly, this relates to the issue of guidance versus exploration in SG-ML, which we have discussed in prior work [22], [23]. In interactive machine learning systems, it is common to require a constant amount of input or interaction from the human partner. A confirmation of our observations in more thorough future studies would indicate that this method is not appropriate for robots learning from everyday human partners. Such robots need a more fluid balance of control of when to lead and when to follow. Thus, an important problem for

future work is designing regulatory mechanisms for switching between leading and following in a social learning interaction.

REFERENCES

- [1] E. M. C. L. S. Vygotsky, *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press, 1978.
- [2] R. M. Krauss, Y. Chen, and P. Chawla, "Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us?" in *Advances in experimental social psychology*, M. Zanna, Ed. Tampa: Academic Press, 1996, pp. 389–450.
- [3] M. Argyle, R. Ingham, and M. McCallin, "The different functions of gaze," *Semiotica*, pp. 19–32, 1973.
- [4] M. Bratman, "Shared cooperative activity," *The Philosophical Review*, vol. 101, no. 2, pp. 327–341, 1992.
- [5] P. R. Cohen, H. J. Levesque, J. H. T. Nunes, and S. L. Oviatt, "Task-oriented dialogue as a consequence of joint activity," in *Proceedings of the Pacific Rim International Conference on Artificial Intelligence*, Nagoya, Japan, November 1990.
- [6] E. Horvitz, J. Breese, D. Heckerman, D. Hovel, and K. Rommelse, "The lumiere project: Bayesian user modeling for inferring the goals and needs of software users," in *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, Madison, WI, July 1998, pp. 256–265.
- [7] Y. Kuniyoshi, M. Inaba, and H. Inoue, "Learning by watching: Extracting reusable task knowledge from visual observation of human performance," *IEEE Transactions on Robotics and Automation*, vol. 10, pp. 799–822, 1994.
- [8] P. Abbeel and A. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Proceedings of the 21st international conference on Machine learning (ICML 2004)*, 2004.
- [9] M. N. Nicolescu and M. J. Mataric, "Natural methods for robot task learning: Instructive demonstrations, generalization and practice," in *Proceedings of the 2nd Intl. Conf. AAMAS*, Melbourne, Australia, July 2003.
- [10] C. Breazeal, A. Brooks, J. Gray, G. Hoffman, J. Lieberman, H. Lee, A. L. (Thomaz), and D. Mulanda, "Tutelage and collaboration for humanoid robots," *International Journal of Humanoid Robotics*, vol. 1, no. 2, 2004.
- [11] R. Maclin, J. Shavlik, L. Torrey, T. Walker, and E. Wild, "Giving advice about preferred actions to reinforcement learners via knowledge-based kernel regression," in *Proceedings of the The Twentieth National Conference on Artificial Intelligence (AAAI)*, Pittsburgh, PA, July 2005.
- [12] A. L. Thomaz and C. Breazeal, "Teachable robots: Understanding human teaching behavior to build more effective robot learners," *Artificial Intelligence Journal*, vol. 172, pp. 716–737, 2008.
- [13] W. D. Smart and L. P. Kaelbling, "Effective reinforcement learning for mobile robots," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2002, pp. 3404–3410.
- [14] S. A. Goldman and H. D. Mathias, "Teaching a smarter learner," *Journal of Computer and System Sciences*, vol. 52, no. 2, pp. 255–267, 1996.
- [15] D. Cohn, Z. Ghahramani, and M. Jordan, "Active learning with statistical models," in *Advances in Neural Information Processing*, G. Tesauro, D. Touretzky, and J. Alsppector, Eds. Morgan Kaufmann, 1995, vol. 7.
- [16] D. Cohn, R. Caruana, and A. McCallum, "Semi-supervised clustering with user feedback," 2003.
- [17] B. Settles, "Active learning literature survey," University of Wisconsin–Madison, Computer Sciences Technical Report 1648, 2009.
- [18] S. Chernova and M. Veloso, "Confidence-based policy learning from demonstration using gaussian mixture models," in *Proc. of Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2007.
- [19] B. Blumberg, M. Downie, Y. Ivanov, M. Berlin, M. Johnson, and B. Tomlinson, "Integrated learning for interactive synthetic characters," in *Proceedings of the ACM SIGGRAPH*, 2002.
- [20] T. M. Mitchell, *Machine Learning*. McGraw-Hill, 1997.
- [21] H. Seung, M. Oppen, and H. Sompolinsky, "Query by committee," in *ACM Workshop on Computational Learning Theory*, 1992, pp. pages 287–294.
- [22] A. L. Thomaz and C. Breazeal, "Reinforcement learning with human teachers: Evidence of feedback and guidance with implications for learning performance," in *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI)*, 2006.
- [23] C. Breazeal and A. L. Thomaz, "Learning from human teachers with socially guided exploration," in *Proceedings of the International Conference on Robots and Automation (ICRA)*, 2008.